Backscattering of Dirac fermions in HgTe quantum wells with a finite gap

G. Tkachov, C. Thienel, V. Pinneker, B. Büttner, C. Brüne, H. Buhmann, L. W. Molenkamp, and E. M. Hankiewicz

Fakultät für Physik und Astronomie and Röntgen Center for Complex Material Systems,
Universität Würzburg, Am Hubland, 97074 Würzburg, Germany

(Dated: April 22, 2011)

The density-dependent mobility of n-type HgTe quantum wells with inverted band ordering has been studied both experimentally and theoretically. While semiconductor heterostructures with a parabolic dispersion exhibit an increase in mobility with carrier density, high quality HgTe quantum wells exhibit a distinct mobility maximum. We show that this mobility anomaly is due to backscattering of Dirac fermions from random fluctuations of the band gap (Dirac mass). Our findings open new avenues for the study of Dirac fermion transport with finite and random mass, which so far has been hard to access.

PACS numbers:

Introduction.– Topological insulators, such as HgTe quantum wells (QWs) [1–3], $Bi_{1-x}Sb_x$ [4, 5], Bi_2Se_3 [6], and Bi_2Te_3 [7, 8] provide the most recent examples of the realization of Dirac fermions in condensed matter physics. Unlike graphene [9], where two valleys of Dirac fermions exist, in these new materials Dirac fermions appear only at a single point in the Brillouin zone. The absence of valley scattering and the possibility of unconventional types of disorder, such as a random Dirac mass [10], make electron transport studies on these materials particularly interesting. However, the low carrier mobility of most topological insulators still poses a serious obstacle for transport investigations of specific scattering mechanisms. The notable exception are MBE-grown HgTe QWs where transport measurements have been used to detect the quantum spin Hall state [2] and, more recently, to realize massless single-valley two-dimensional Dirac fermions [11].

In this paper, we study both experimentally and theoretically a new manifestation of Dirac fermion transport in n-type HgTe QWs, i.e., the occurrence of a maximum in the mobility as a function of carrier density. The origin of the maximum is the competition of two disorder scattering mechanisms, viz. scattering by charged impurities and by QW width fluctuations which induce a fluctuating band gap, or, equivalently, fluctuating Dirac mass. As in other semiconductor heterostructures [12], in HgTe QWs the screening of ionized impurities by the carriers results, initially, in a monotonic increase of the mobility with increasing carrier density. However, while the impurity scattering becomes weaker with increasing carrier density, the other scattering mechanism - well-width fluctuations - becomes increasingly important, leading to a reduction of the carrier mobility. Dirac mass disorder generates scattering between states of opposite momenta, also called backscattering. Thus, the observed mobility peak is a clear manifestation of Dirac fermion backscattering in HgTe QWs.

Backscattering of Dirac fermions is most pronounced in gapped systems [13]. As a consequence, the mobil-

ity of graphene does not show a maximum, but rather a saturation at high carrier densities (see, e.g. Refs. [9, 13–15]), which has been attributed to charged impurity scattering. Although some theoretical studies indicate that mass disorder may play a role in the vicinity of the neutrality point of graphene (see, e.g. Refs. [16, 17]), its experimental identification has remained problematic.

Experiment.– Transport experiments have been performed on modulation doped $HgTe/Hg_{0.3}Cd_{0.7}Te$ QW structures fabricated by molecular beam epitaxy on lattice-matched (Cd,Zn)Te substrates. The samples have been patterned into Hall bar devices with dimensions of (600×200) μm^2 using a low temperature optical lithography process, and covered by a 5/100 nm Ti/Au gate electrode which is deposited onto a 110 nm thick Si_3N_4/SiO_2 multilayer gate insulator. Ohmic contacts are provided by thermal indium bonding. A micrograph of the structure is shown in the inset of Fig. 1. The samples have nominal QW widths, d, ranging from 5.0 to 12.0 nm, thus covering both the normal ($d < 6.3$ nm) and inverted ($d > 6.3$ nm) band structure regimes [2, 11]. The relevant
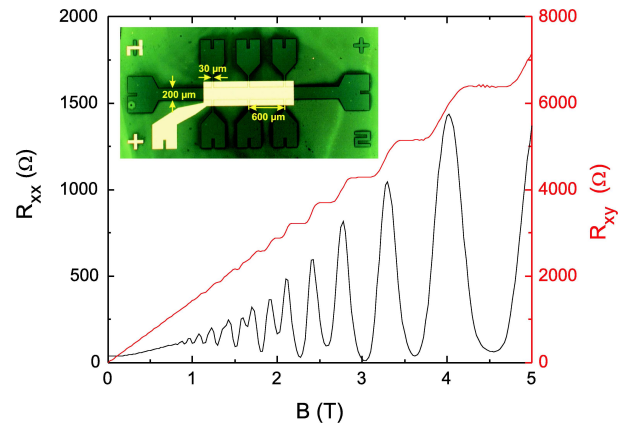

FIG. 1: Hall and magnetoresistance for sample #6 at zero gate voltage. The measurement yields an electron mobility of 1.1×10^6 cm^{-2}/Vs and a carrier density of 4.3×10^{11} cm^{-2}. Inset: Micrograph of the Hallbar structure.

parameters of all samples are summarized in Table I. We have performed standard Hall and Shubnikov-de Haas measurements on these samples in magnetic fields up to $B = 5$ T, at a temperature 4.2 K. As an example, Hall and magnetoresistance data for sample #6, which has the peak highest mobility are shown in Fig. 1. The carrier densities and mobility of ungated samples ($V_g = 0$) are in the range of 3.0×10^{11} cm^{-2} to 5.5×10^{11} cm^{-2} and several 10^5 cm^2/Vs, respectively.

The density dependent carrier mobilities are obtained from the gate-voltage dependence of the longitudinal resistance at zero magnetic field, while the dependence of the carrier concentration on gate-voltage was deduced from the Hall voltage measured at a fixed magnetic field of 300 mT. The thick solid lines in Fig. 2 show the experimentally observed mobility versus carrier density for the various HgTe QWs. In high quality samples with inverted subband structure ordering (#3, #4, and #6) the mobility exhibits a distinct maximum for carrier densities in the range of 3 to 6 $\times 10^{11}$ cm^{-2}. For samples with a slightly lower mobility (#1, #2, and #5) a saturation of the mobility with carrier density is observed.

Model.– In order to explain the unusual dependence of the mobility on carrier density we build a model for the conductivity in HgTe QWs which is based on the four-band Dirac model of Refs. [1, 2]. The effective Dirac Hamiltonian has the following form:

$$\hat{H} = s_z\boldsymbol{\sigma}(A\hat{\mathbf{k}} + M_{\hat{\mathbf{k}}}\mathbf{z}) + D\hat{\mathbf{k}}^2 - E_F + V_{\mathbf{r}} + \delta M_{\mathbf{r}} s_z \sigma_z, \quad (1)$$

where $M_{\hat{\mathbf{k}}} = M(d) + B\hat{\mathbf{k}}^2$. In Eq. (1) the linear term, which is proportional to $\hat{\mathbf{k}} = -i\nabla_{\mathbf{r}}$ and the constant A originate from the hybridization of the first electron (E1) and heavy-hole (HH1) subbands in the quantum well. These two subbands are represented by the pseudospin $\boldsymbol{\sigma}$, whose components σ_x, σ_y and σ_z are 2×2 Pauli matrices. (The real spin degree of freedom is represented by the Pauli matrix s_z.) The term proportional to the effective Dirac mass $M_{\hat{\mathbf{k}}}$ reflects the average band gap, $|M_{\hat{\mathbf{k}}=\mathbf{0}}| = |M(d)|$, which is determined by the nominal thickness d of the QW. The $\mathbf{k}$-dependent part of the mass (the $B\hat{\mathbf{k}}^2$-term) and the parabolic background $D\hat{\mathbf{k}}^2$ take further details of the band dispersion in HgTe QWs into account [1, 2]. Note that the mass term violates the pseudo-time reversal symmetry $\hat{\mathbf{k}} \to -\hat{\mathbf{k}}$ and $\boldsymbol{\sigma} \to -\boldsymbol{\sigma}$ of Hamiltonian (1), which manifests itself in a dependence of the conductivity on $M_{\hat{\mathbf{k}}}$. Equation (1) also takes the two most relevant types of disorder into account: a random potential due to charged impurities, $V_{\mathbf{r}}$, and spatial fluctuations of the Dirac mass, $\delta M_{\mathbf{r}}$, which are related to deviations of the QW thickness from its average value d. We assume the n-type transport regime, with the Fermi level E_F in the conduction band, under the weak scattering condition, $k_F v_F \tau \gg 1$, where τ is the transport relaxation time and v_F and $k_F = \sqrt{2\pi n}$ are the Fermi velocity and wave-vector, respectively. The conductivity
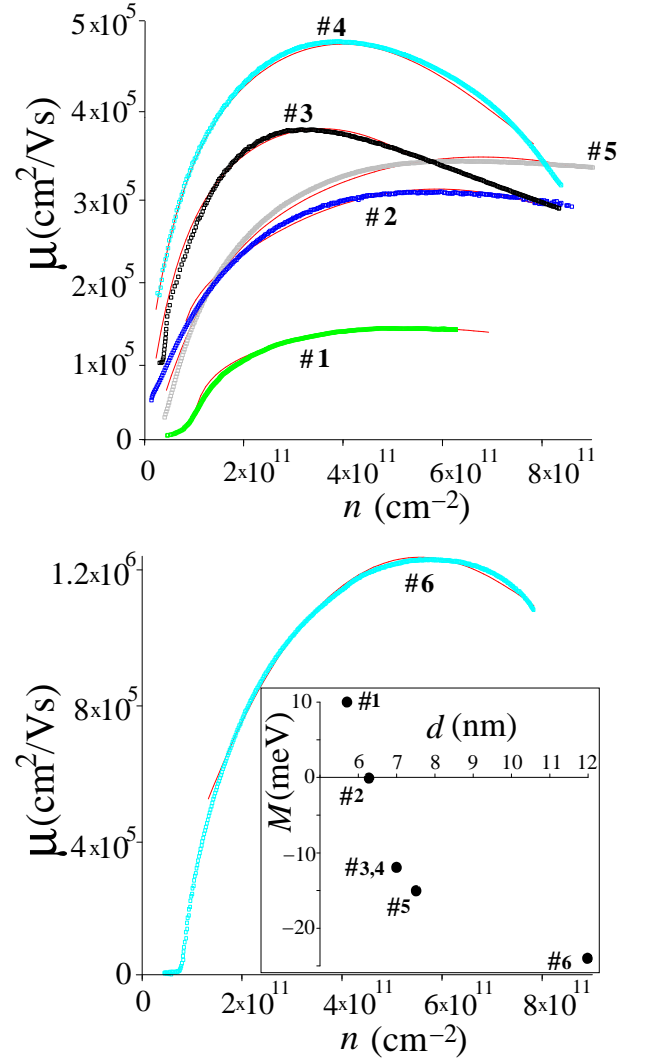

FIG. 2: Mobility μ versus carrier density n for six HgTe QWs (#1 − #6). The experimental $\mu(n)$ dependence (thick lines), including the appearance of the maximum, agrees well with our model (thin red lines) that takes into account the competition of charged-impurity disorder and QW thickness fluctuations. Inset: Band gap M versus QW thickness d [sample #1 has a positive band gap, sample #2 has approximately zero gap, while the rest #3 − #6 are in the inverted regime]. The pronounced maximum is observed in high quality inverted samples #3, #4 and #6 where the impurity concentration is sufficiently low (see also Table I).

can then be obtained from the Kubo formula:

$$\sigma_{xx} = \frac{e^2\hbar}{2\pi} \int \frac{d^2\mathbf{k}}{(2\pi)^2} \text{Tr}\left[\hat{v}_x \hat{G}_{\mathbf{k}}^R \hat{\tilde{v}}_x \hat{G}_{\mathbf{k}}^A\right] = e^2 N_F v_F^2 \tau, \quad (2)$$

where $\hat{G}_{\mathbf{k},\epsilon}^{R/A}$ are disorder-averaged retarded and advanced Green's functions and $\hat{v}_x$, $\hat{\tilde{v}}_x$ are current vertices (the tilde indicates vertex renormalization by disorder in the ladder approximation [18]). The resulting conductivity σ_{xx} is proportional to the density of states per spin at the Fermi level, $N_F = k_F/2\pi\hbar v_F$, and the transport time

$\tau = 1/\int_0^{2\pi}(1 - \cos\theta)w(\theta)d\theta$, where $w(\theta)$ is the scattering rate at angle θ. Below, we obtain an expression for $w(\theta)$ from the electron self-energy in the self-consistent Born approximation (SCBA), calculate τ and the carrier-density dependent mobility $\mu(n) = \sigma_{xx}(n)/ne$.

For this purpose, we assume that the potential and mass disorder are uncorrelated and completely characterized by the two-point correlation functions $\langle V_{\mathbf{r}}V_{\mathbf{r}'}\rangle = \zeta^V_{\mathbf{r}-\mathbf{r}'}$ and $\langle \delta M_{\mathbf{r}}\delta M_{\mathbf{r}'}\rangle = \zeta^M_{\mathbf{r}-\mathbf{r}'}$. In $\mathbf{k}$ space this leads to the Dyson equation $\hat{G}^R_{\mathbf{k}} = \hat{G}^R_{0\mathbf{k}} + \hat{G}^R_{0\mathbf{k}}\hat{\Sigma}^R_{\mathbf{k}}\hat{G}^R_{\mathbf{k}}$ for the disorder-averaged Green's function $\hat{G}^R_{\mathbf{k}}$ where the self-energy $\hat{\Sigma}^R_{\mathbf{k}}$ is given by the standard SCBA expression

$$\hat{\Sigma}^R_{\mathbf{k}} = \int (\zeta^V_{\mathbf{k}-\mathbf{q}}\hat{G}^R_{\mathbf{q}} + \zeta^M_{\mathbf{k}-\mathbf{q}}s_z\sigma_z\hat{G}^R_{\mathbf{q}}s_z\sigma_z)\frac{d\mathbf{q}}{(2\pi)^2}, \quad (3)$$

$$\hat{G}^R_{0\mathbf{k}} = \frac{1}{2}\frac{I + s_z\boldsymbol{\sigma}\cdot\mathbf{e_k}}{\epsilon - \xi_{\mathbf{k}}}, \quad \mathbf{e_k} = \frac{A\mathbf{k} + M_{\mathbf{k}}\mathbf{z}}{\sqrt{A^2\mathbf{k}^2 + M^2_{\mathbf{k}}}}. \quad (4)$$

The unperturbed Green's function $\hat{G}^R_{0\mathbf{k}}$ describes a conduction-band electron with dispersion $\xi_{\mathbf{k}} = \sqrt{A^2\mathbf{k}^2 + M^2_{\mathbf{k}}} + D\mathbf{k}^2 - E_F$ and chirality $s_z\boldsymbol{\sigma}\cdot\mathbf{e_k} = 1$ (I is the unit matrix). The solution [18] for the Green's function $\hat{G}^R_{\mathbf{k}} = \frac{1}{2}(I + s_z\boldsymbol{\sigma}\cdot\mathbf{e_k})/(\epsilon - \xi_{\mathbf{k}} + \frac{i\hbar}{2\tau_{el}})$ contains the finite elastic life-time $\tau_{el} = 1/\int_0^{2\pi}w(\theta)d\theta$, where the scattering rate $w(\theta)$ at angle θ is given by

$$w(\theta) = \frac{N_F}{\hbar}\left[\left(\zeta^V_{2k_F|\sin\theta/2|} + \zeta^M_{2k_F|\sin\theta/2|}\mathrm{e}^2_{\perp}\right)\cos^2\frac{\theta}{2}\right.$$

$$\left. + \left(\zeta^V_{2k_F|\sin\theta/2|}\mathrm{e}^2_{\perp} + \zeta^M_{2k_F|\sin\theta/2|}\right)\sin^2\frac{\theta}{2}\right]. \quad (5)$$

Here $\mathrm{e}_{\perp} = M_{k_F}/\sqrt{A^2k^2_F + M^2_{k_F}}$ is the out-of-plane component of the unit vector $\mathbf{e_k}$ at $|\mathbf{k}| = k_F$ [see Eq. (4)].

Analyzing Eq. (5), one notes that the first term vanishes at $\theta = \pi$. This is the well-known absence of backscattering in the limit of massless Dirac electrons [13] (this behavior is plotted as the red curve in Fig. 3a). In gapless Dirac materials, the pseudospin $\boldsymbol{\sigma}$ points along $\mathbf{k}$, and, therefore, states with $\mathbf{k}$ and $-\mathbf{k}$ are orthogonal to each other, i.e. unavailable for scattering. The most essential distinction between HgTe quantum wells and the zero gap case is the second term in Eq. (5), which actually has a maximum at the backscattering angle $\theta = \pi$. This term originates from the finite Dirac mass $M_{\mathbf{k}}s_z\sigma_z$ and its spatial fluctuation $\delta M_{\mathbf{r}}s_z\sigma_z$. Both of these result in an out-of-plane pseudospin component σ_z, hence the opposite $\mathbf{k}$ states are no longer orthogonal and large-angle scattering now becomes possible (cf. the black curve in Fig. 3a). Figure 3b shows that the backscattering is enhanced with increasing carrier density n, which accounts for the non-monotonic behavior of $\mu(n)$ in inverted quantum wells observed in Fig. 2.

To proceed further, we make specific assumptions for the correlation functions:

$$\zeta^V_{\mathbf{k}} = n_i(2\pi e^2)^2/(\varepsilon|\mathbf{k}| + 4\pi e^2 N_F)^2, \quad \zeta^M_{\mathbf{k}} = A^2\zeta^M, \quad (6)$$
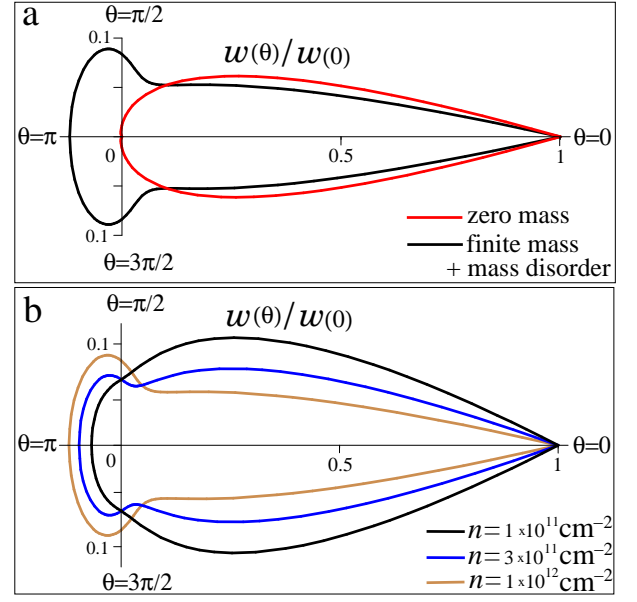

FIG. 3: (a) Polar plot of the angular dependence of the normalized scattering rate $w(\theta)$ for massless and massive ($M = -15$ meV and $\zeta^M = 1.5 \times 10^{-3}$) Dirac fermions [see Eqs. (5) and (6)]. Massive carriers show strong backscattering at angles $\pi/2 < \theta < 3\pi/2$. (b) Polar plot of the normalized function $w(\theta)$ for different carrier densities n in the inverted regime ($M = -15$ meV and $\zeta^M = 1.5 \times 10^{-3}$). The large-angle scattering is enhanced with increasing n, which accounts for the non-monotonic mobility $\mu(n)$ in Fig. 2.

where $\zeta^V_{\mathbf{k}}$ is the usual correlation function of screened Coulomb impurities [13, 14] with the concentration n_i and the average dielectric constant, $\varepsilon \approx 15$, of the HgTe/CdTe QW. $\zeta^M_{\mathbf{k}}$ is normalized such that ζ^M is a small dimensionless parameter, which guarantees that δM is small compared with the leading linear term $\sim Ak_F$ in Hamiltonian (1). Since δM is caused by fluctuations of the QW thickness, ζ^M is independent of carrier density. As shown below, this approximation yields an excellent agreement with the measurements. Using Eqs. (2), (5) and (6) we find the mobility $\mu(n) = \sigma_{xx}/ne$:

$$\mu(n) = \frac{4e/h}{n_i[\zeta^V_1(n) + \zeta^V_2(n)\mathrm{e}^2_{\perp}(n)] + n\zeta^M\frac{[3+\mathrm{e}^2_{\perp}(n)]A^2}{\hbar^2 v^2_F(n)}}, \quad (7)$$

where

$$\mathrm{e}_{\perp}(n) = (M + 2\pi Bn)/\sqrt{2\pi A^2 n + (M + 2\pi Bn)^2}, \quad (8)$$

$$\zeta^V_1 = \int_0^{\pi} d\theta \sin^2\theta \left[1 + \varepsilon\hbar v_F(n)e^{-2}\sin\theta/2\right]^{-2}, \quad (9)$$

$$\zeta^V_2 = 4\int_0^{\pi} d\theta \sin^4\theta/2 \left[1 + \varepsilon\hbar v_F(n)e^{-2}\sin\theta/2\right]^{-2}. \quad (10)$$

Using Eq. (7) we can quantitatively reproduce all experimental curves for $\mu(n)$ in Fig. 2 using the disorder parameters n_i, ζ^M and band gap M indicated in Table I. The values of M obtained from this fit agree well with

TABLE I: Sample and disorder parameters (the other parameters of the model are fixed: $\varepsilon = 15$, $A = 0.38$ eV·nm, $B = 1.2$ eV·nm^2 and $D = 0.85$ eV·nm^2).

	# 1	# 2	# 3	# 4	# 5	# 6
nominal well width (nm)	5.7	6.3	7	7	7.5	12
max. mobility (10^5 cm^2/Vs)	1.32	2.96	3.71	4.76	3.33	12.27
M (meV)	10	0	-12	-12	-15	-24
$n_i(\times 10^{10}$cm$^{-2})$	8	3.41	2.95	2.55	4.56	1.09
$\zeta^M(\times 10^{-3})$	1.65	0.5	1.65	0.91	0.43	0.33

those obtained from band structure calculations and the analysis of the experimental SdH oscillations.

Discussion.– Using the theoretical model presented in the previous section, we can now explain the observed non-monotonic dependence of $\mu(n)$ as follows. The initial increase in $\mu(n)$ results from the impurity screening: it reflects the density-dependence of the Fermi velocity $v_F(n)$ which enters the screened impurity potential via the DOS in Eq. (6) [see also Eqs. (9) and (10)]. Similar behavior was found in conventional doped heterostructures [12]. However, here $v_F(n)$ is specific to the massive Dirac Hamiltonian (1). Furthermore, for the inverted quantum wells $\#3 - \#6$, the mobility is additionally enhanced due to the reduction of the total Dirac gap $-|M(d)| + 2\pi Bn$ at low carrier densities n [cf. Eq. (8)]. This leads to a more rapid initial increase in $\mu(n)$ compared to sample $\#1$, which has a normal band structure, and the zero-gap sample $\#2$. At higher carrier densities the mobility starts to decrease for all the inverted samples, most pronouncedly so for the high-quality samples $\#3$, $\#4$ and $\#6$. Since the estimated impurity concentration is lowest for these samples (see Table I), we attribute this decrease to the fluctuations of the QW width (Dirac mass), accounted for by the term $\propto n\zeta^M$ in Eq. (7). Intuitively, the reduction of the mobility can be explained by the fact that the rate of scattering off the well-width fluctuations grows proportionally to the carrier DOS $N_F(n)$ because more states become available for backscattering as the Fermi surface size increases with n [see Eqs. (5) and (6)].

From the fits we estimate the amplitude of the well width fluctuations, δd, to be of the order of $0.2 - 0.3$ nm, which is obtained by integrating the correlation function of the thickness fluctuations, $\langle \delta d_{\mathbf{r}} \delta d_{\mathbf{r}'} \rangle = \langle \delta M_{\mathbf{r}} \delta M_{\mathbf{r}'} \rangle / M'(d)^2$, over the area. This integral is a measure of the typical height δd times length L of the fluctuation: $L\delta d \sim A\sqrt{\zeta^M}/|M'(d)| \approx 0.7$ nm^2, where we take $A = 0.38$ eV·nm and $\zeta^M = 10^{-3}$ from Table I and the gap derivative $|M'(d)| \approx 17$ meV·nm^{-1} from the inset in Fig. 2. For a realistic sample $L \gg \delta d$, thus we estimated the ratio $L/\delta d \approx 10$, which yields $\delta d \approx 0.26$ nm. The result is in good agreement with X-ray reflectivity data on MBE-grown HgTe QW structures of similar quality [19].

Conclusions.– We have shown both experimentally and

theoretically that the density-dependent mobility of high quality HgTe quantum wells with an inverted band structure exhibits a maximum. While the initial increase in the mobility is mainly due to scattering at screened charged impurities, the decreasing part is associated with the band gap fluctuations that generate mass disorder for Dirac-like fermions in this material. Our findings thus clearly demonstrate the occurrence of Dirac fermion backscattering in finite gap systems.

[1] B. A. Bernevig and T. L. Hughes and S. C. Zhang, Science **314**, 1757 (2006).
[2] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi and S.-C. Zhang, Science **318**, 766 (2007).
[3] A. Roth, C. Brüne, H. Buhmann, L. W. Molenkamp, J. Maciejko, X.-L. Qi, and S.-C. Zhang, Science **325**, 294 (2009).
[4] L. Fu and C. L. Kane, Phys. Rev. B **76**, 045302 (2007).
[5] D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nature Phys. **452**, 970974 (2008).
[6] Y. Xia, D. Qian, D. Hsieh, L. Wray, A. Pal, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nature Phys. **5**, 398 (2009).
[7] Y. L. Chen, J. G. Analytis, J.-H. Chu, Z. K. Liu, S.-K. Mo, X. L. Qi, H. J. Zhang, D. H. Lu, X. Dai, Z. Fang, S. C. Zhang, I. R. Fisher, Z. Hussain, and Z.-X. Shen, Science **325**, 178 (2009).
[8] H. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang and S.-C. Zhang, Nature Phys. **5**, 438 (2009).
[9] A.H. Castro Neto, F. Guinea, N.M. Peres, K.S. Novoselov, and A.K. Geim, Rev. Mod. Phys. **81**, 109 (2009) and references therein.
[10] A. W. W. Ludwig, M. P. A. Fisher, R. Shankar, and G. Grinstein, Phys. Rev. B **50**, 7526 (1994).
[11] B. Büttner, C. X. Liu, G. Tkachov, E. G. Novik, C. Brüne, H. Buhmann, E. M. Hankiewicz, P. Recher, B. Trauzettel, S. C. Zhang and L. W. Molenkamp, Nature Phys. (published online 6 February 2011, doi:10.1038/nphys1914); e-print arXiv:1009.2248.
[12] M. S. Shur, J. K. Abrokwah, R. R. Daniels, and D. K. Arch, J. Appl. Phys **61**, 1643 (1987).
[13] T. Ando, J. Phys. Soc. Jpn. **74**, 777 (2005).
[14] K. Nomura and A.H. MacDonald, Phys. Rev. Lett. **96**, 256602 (2006).
[15] J.-H. Chen, C. Jang, S. Adam, M. S. Fuhrer, E. D. Williams and M. Ishigami, Nature Phys. **4**, 377 (2008).
[16] K. Ziegler, Phys. Rev. Lett. **102**, 126802 (2009).
[17] J. H. Bardarson, M. V. Medvedyeva, J. Tworzydlo, A. R. Akhmerov, and C. W. J. Beenakker, Phys. Rev. B **81**, 121414(R) (2010).
[18] J. Rammer, *Quantum Transport Theory* (Westview Press, 2004).

[19] A. Stahl, Ph. D. Thesis, University of Würzburg, 2010.